4-Ч

УДК 004.032.26:004.8

П.В. Четырбок
Ялтинский филиал Европейского университета, г. Ялта, Украина
petr58@mail.ru

Алгоритм сжатия информации с использованием скалярного критерия распознавания образов

В данной работе построено изоморфное отображение множества образов на множество вещественных чисел из интервала [0,1], которое позволяет связать классификацию образов с поведением скалярного критерия в пространстве ошибок. Задавая определенные интервалы для значений скалярного критерия, можем группировать образы, распознавать, сравнивать и анализировать их. В работе построен алгоритм сжатия информации с использованием скалярного критерия.

Постановка проблемы

Предложенный в статье алгоритм сжатия информации основан на использовании скалярного критерия для распознавания образов нейронной сетью. На вход нейронной сети подается двоичный ASCII код информации и рассчитывается скалярный критерий для его распознавания. Так для кодирования сообщения «Алгоритм сжатия!» в ASCII коде необходимо 128 битов, а скалярный критерий, являясь действительным числом двойной точности требует 64 бита. То есть получаем сжатие информации в 2 раза. Но для этого необходимо построить многослойную нейронную сеть с 128 нейронами на входе и 128 выходными нейронами, которая является классификатором двоичной информации на входе. То есть имеем задачу классификации.

Для задачи классификации M классов, в которой объединение M классов формирует все пространство входных образов, для представления всех возможных результатов классификации требуется M выходов [1]. Каким должно быть оптимальное решающее правило, применяемое для классификации M выходов сети после обучения многослойного персептрона? В общем случае решающее правило должно основываться на знании вектор-функции [1] $F: R^M x \to y \in R^M$. Обычно в качестве вектор-функции берется непрерывная функция, минимизирующая функционал эмпирического риска [1]:

$$R = \frac{1}{2N}\sum_{j=1}^{N}\left\| d_j - F(x_j) \right\|^2 ,$$

где d_j – желаемый выход для прототипа x_j; $\| \; \|$ – Евклидова норма вектора; N – общее число примеров, представленных сети для обучения. Построить функционал и оптимальное решающее правило. В статье построен функционал, равный скалярному произведению векторов ошибок при распознавании нейронной сетью образов и соответствующих им эталонов, который является вектор-функцией.

Цель статьи

Построение функционала и решающего правила (скалярного критерия для распознавания образов) для решения задачи сжатия информации. Функционал равен скалярному произведению векторов ошибок при распознавании нейронной сетью образов и соответствующих им эталонов. Впервые построено решающее правило для классификации образов в виде утверждения: каждому образу, распознаваемому многослойным персептроном в многофакторном пространстве ошибок будет соответствовать свое значение *cos(λ)* и образ ближе к эталону, чем больше *cos(λ*).

$$\cos(\lambda)=\frac{(\overline{E},\overline{X})}{\left\|\overline{E}\right\|_c\times\left\|\overline{X}\right\|_c},$$

где E – вектор ошибок в пространстве ошибок, полученный при распознавании нейронной сетью входного образа, X – вектор ошибок, полученный при распознавании нейронной сетью эталона. Что позволяет связать классификацию распознаваемых образов с поведением *cos(λ)* в многофакторном пространстве ошибок.

Формулировка скалярного критерия близости образов в пространстве ошибок

Переход из пространства параметров в пространство ошибок. Для распознавания трудноразличимых образов построим функционал

$$F = E_2\times x_1+E_3\times x_2+E_4\times x_3, \qquad (1)$$

E_2 – среднеквадратическая ошибка в пространстве параметров образа, полученная при обучении сети; $E_2=\frac{1}{2}\sum_{j,p}(y_{j,p}^{(N)}-d_{j,p})^2$, где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N нейронной сети при подаче на ее входы p-го образа; d_{jp} – идеальное (желаемое) выходное состояние этого нейрона.

E_3 – линейная ошибка сети, полученная как сумма модулей параметрического отклонения образа от эталона;

$$E_3=\frac{1}{2}\sum_{j,p}\left|y_{j,p}^{(N)}-d_{j,p}\right|,$$

где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N нейронной сети при подаче на ее входы p-го образа; d_{jp} – идеальное (желаемое) выходное состояние этого нейрона.

E_4 – максимальная ошибка параметрического отклонения образа от эталона.

$$E_4=\frac{1}{2}\sum_{j,p}\max_{j,p}\left|y_{j,p}^{(N)}-d_{j,p}\right|,$$

где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N нейронной сети при подаче на ее входы p-го образа; d_{jp} – идеальное (желаемое) выходное состояние этого нейрона.

x_1, x_2, x_3 – коэффициенты функционала, полученные при обучении сети. x_1 – среднеквадратическая ошибка, полученная при обучении сети, при заданном эталоне на входе сети; x_2 – линейная ошибка сети, полученная как сумма модулей параметрического отклонения образа от эталона, при заданном эталоне на входе сети; x_3 – максимальная ошибка параметрического отклонения образа от эталона, при заданном эталоне на входе сети.

Для каждого из распознаваемых образов получим свое значение функционала, соответствующее этому образу при обучении сети. Если даже для двух различных образов величины E_2, E_3, E_4 совпадут, значения функционала для этих образов будут различны. Если в качестве образов использовать волны, то первый член функционала отвечает за отклонение частоты образа от эталонной, второй член отвечает за отклонение по амплитуде волны образа от эталонной, третий член анализирует пиковые всплески сигнала. Функционал имеет нижнюю границу при $F = 0$.

Для образов волновой природы функционал позволяет выделить вредные гармоники, подавить шумы и выделить полезный сигнал.

В соответствии с введенным функционалом (1), минимизируемой целевой функцией ошибки НС является величина:

$$E(w) = \frac{1}{2}\sum_{j,p}(y_{j,p}^{(N)} - d_{j,p})^2, \quad (2)$$

где $y_{j,p}^{(N)}$ – реальное выходное состояние нейрона j выходного слоя N нейронной сети при подаче на ее входы p-го образа; d_{jp} – идеальное (желаемое) выходное состояние этого нейрона.

Суммирование ведется по всем нейронам выходного слоя и по всем обрабатываемым сетью образам. Минимизация ведется методом градиентного спуска [1].

При распознавании образа нейронной сетью вычисляем функционал, используя весовые коэффициенты, полученные при обучении сети.

Введем взвешенную норму в сепарабельном трехмерном пространстве ошибок распознавания образов (3).

$$\|X\| = \sqrt{(CX)^T(CX)}, \quad (3)$$

где C – матрица взвешенной нормы.

Функционал F, полученный при обучении сети, равен скалярному произведению векторов ошибок (E2,E3,E4) и (×1, ×2, ×3).

$$F = (\vec{E}, \vec{X}) = \cos(\lambda) * \|\vec{E}\| \times \|\vec{X}\|. \quad (4)$$

Из определения нормы имеем:

$$\|\overline{E}\| = \sqrt{(CE)^T(CE)},$$

$$\|\overline{X}\| = \sqrt{(CX)^T(CX)}.$$

Найдем косинус угла между векторами ошибок:

$$\cos(\lambda) = \frac{(\overline{E}, \overline{X})}{\left\|\overline{E}\right\| \times \left\|\overline{X}\right\|}. \qquad (5)$$

Аналогично находим косинусы для всех эталонных образов F_1, F_2,…, F_n , где n – число образов, предъявленных сети при обучении, и сравниваем с *cos(λ)*, полученным при распознавании нейронной сетью входного образа. Тот функционал F_i, для которого *cos(λ)* наиболее близок к единице, определяет соответствующий входному образу эталонный образ. Взвешенный критерий близости образов в пространстве ошибок (5) получен по аналогии с теоремой равновесия в задачах линейного программирования, которая утверждает, что скалярное произведение вектора решения прямой задачи на вектор решения обратной задачи равно нулю. То есть $(E, X) = 0$, где E – вектор ошибок в пространстве ошибок, полученный при распознавании нейронной сетью входного образа, а X – вектор ошибок, полученный при распознавании нейронной сетью эталона. Вектор Е – вектор решения прямой задачи линейного программирования, а X – вектор решения обратной задачи линейного программирования. Для расчета взвешенного критерия используется взвешенная норма в пространстве ошибок. В общем случае вектора ошибок $E = (e1, e2, \ldots, e_k)$ и $X = (x1, x2, \ldots, x_k)$, где k – количество факторов, влияющих на ошибку распознавания образов многослойным персептроном, т.е. это не размерность пространства образов. Входной образ $Y_{вх} = (y1, y2, \ldots, y_m)$. Эталонный образ $D_{вых} = (d1, d2, \ldots, d_m)$, где m – количество входных нейронов, равное количеству выходных нейронов в многослойном персептроне. Вектор E получен при распознавании многослойным персептроном входного образа $Y_{вх}$ за эталон был взят $D_{вых}$, а вектор X получен, когда на вход сети подали $D_{вых}$ и ожидали $Y_{вх}$. Теперь можно оценить близость распознаваемых образов к эталону по правилу (взвешенный критерий): образ ближе к эталону, чем больше *cos(λ)*.

$$\cos(\lambda) = \frac{(\overline{E}, \overline{X})}{\left\|\overline{E}\right\|_C \times \left\|\overline{X}\right\|_C}.$$

В случае когда образ совпадает с эталоном $cos(\lambda) = 1$ (не учитывая вычислительные ошибки алгоритма).

Выводы

Каждому образу, распознаваемому многослойным персептроном в многофакторном пространстве ошибок соответствует свое значение скалярного критерия *cos(λ)*. В данной работе построено отображение множества распознаваемых образов на множество действительных чисел из интервала *[0,1]*, которое позволяет связать классификацию распознаваемых образов с поведением *cos(λ)* в многофакторном пространстве ошибок. Задавая определенные интервалы для значений *cos(λ),* мы можем группировать образы, распознавать, сравнивать и анализировать их. Впервые построено решающее правило для классификации образов в виде утверждения: каждому образу, распознаваемому многослойным персептроном в многофакторном пространстве ошибок, будет соответствовать свое значение *cos(λ)* и образ ближе к эталону, чем больше *cos(λ)*.

$$\cos(\lambda) = \frac{(\overline{E}, \overline{X})}{\left\|\overline{E}\right\|_C \times \left\|\overline{X}\right\|_C},$$

где E – вектор ошибок в пространстве ошибок, полученный при распознавании нейронной сетью входного образа, X – вектор ошибок, полученный при распознавании нейронной сетью эталона. Предложенный в статье алгоритм сжатия информации основан на использовании скалярного критерия для распознавания образов нейронной сетью. На вход нейронной сети подается двоичный ASCII код информации и рассчитывается скалярный критерий для его распознавания.

Литература

1. Хайкин Саймон. Нейронные сети: Полный курс, 2-е издание: Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
2. Жураковський Ю.П., Полторак В.П. Теорія інформації та кодування. – Видавництво «Вища школа», 2002. – 255 с.
3. Горбань А.Н. Обучение нейронных сетей. – М.: Изд. СССР-США СП «ПараГраф», 1990. – 160 с.
4. Кохонен Т. Ассоциативная память. – М.: Мир, 1980.
5. Рассоха А.А., Четырбок П.В. Скалярный критерий близости в пространстве распознаваемых образов. – Европейский университет. – Ялта, 2007. – 26 с.–Укр.-Деп. в ГНТБ Украины
6. Амосов Н.М. Алгоритмы разума. – Киев: наукова думка, 1979. – 222 с.
7. Глибовець М.М., Олецький О.В. Штучний інтелект. – Київ: «КМ Академія», 2000.
8. Richard M.D. and Lipman R.P. Neural network classifiers estimate Bayesian a posteriori probabilities, Neural Computation. 1991. – Vol. 3. – p. 461-483.
9. White H. Learning in artificial neural networks: A statistical perspective, Neural Computation. 1989. – Vol.1. – P. 425-464.

П.В. Четирбок

Алгоритм стиснення інформації з використанням скалярного критерію розпізнавання образів

У даній роботі побудовано ізоморфне відображення безлічі образів на безліч дійсних чисел з інтервалу [0,1], яке дозволяє пов'язати класифікацію образів з поведінкою скалярного критерію в просторі помилок. Задаючи певні інтервали для значень скалярного критерію, можемо групувати образи, розпізнавати, порівнювати і аналізувати їх. У роботі побудований алгоритм стиснення інформації з використанням скалярного критерію.

P.V. Chetyrbok

Algorithm of Information Compression using Scalor Criteria of Image Recognition

In the given work the displaying is built multiplicands of recognizable appearances on multiplied material numbers from interval [0,1], which allows to link classification of recognizable appearances with the conduct of scalar criterion in space of errors. Setting definite intervals for the values of scalar criterion can group appearances, recognize, compare and analyze them. In work algorithm of compression of information with the use of scalar criterion is built.

Статья поступила в редакцию 10.07.2008.